Future Technology
Open Access Journal
ISSN 2832-0379
May 2022| Volume 01 | Issue 01 | Pages 01-06
Journal homepage: https://fupubco.com/futech
https://doi.org/10.55670/fpll.futech.1.1.1
Future Publishing LLC

Review

Application of machine learning in flood forecasting

Mehrsa Bayat[1]*, Omid Tavakkoli[2]

[1]Department of Bioresource Engineering, McGill University, Montreal, Canada
[2]Department of Petroleum Engineering, Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Malaysia

ARTICLE INFO

Article history:
Received 01 March 2022
Received in revised form
26 March 2022
Accepted 31 March 2022

Keywords:
Flood Forecasting, Machine Learning, Deep Learning, Time-Series Prediction

*Corresponding author
Email address: mehrsa.bayat@mail.mcgill.ca

DOI: 10.55670/fpll.futech.1.1.1

ABSTRACT

A flood is a costly natural disaster that imposes a considerable risk to many urban areas worldwide. Predicting the flood can help to alleviate the damage that it causes. In recent years, inspired by the success of Machine Learning (ML) in other fields, several papers have proposed ML-based algorithms for short-term and long-term flood prediction. In this study, we aim to give an overview of ongoing research in this area. We present several case studies from recent papers that employed machine learning for flood forecasting. Results of these studies have shown that ML models are powerful tools in flood prediction. We also briefly reviewed some of the commonly used machine learning models. However, the technical description of these methods is beyond the scope of this study and is not discussed in detail. Although ML models have shown promising results in flood forecasting, they do suffer from important limitations, and there is still room for their improvement. In this paper, we attempted to critically assess the shortcomings of existing algorithms and offer several suggestions for the future direction of this research field.

1. Introduction

1.1 Background

According to World Health Organization (WHO), floods are the most common natural disaster, affecting more than 2 billion people globally between 1998-2017 [1]. In Canada, floods cause damage to properties more than any other natural disaster. Since most Canadian cities are developed along rivers and lakes, floods pose a serious threat to urban areas [2]. Generally, flooding occurs when the volume of water in a channel exceeds its capacity. In Canada, most flooding takes place during spring when the frozen snow or ice starts to melt rapidly. Excessive rain, a ruptured dam, large storms, and tsunamis can also raise the water level and lead to a flood. Every year, flooding leads to substantial financial losses for governments, and many people are involuntarily displaced as a result of this devastating natural phenomenon [3]. Early forecasting of floods can significantly reduce damages and economic losses. Therefore, hydrologists have developed a flurry of models for short-term and long-term prediction of flooding [4]. Conventionally, physically-based methods, designed by incorporating expert knowledge, have been used to predict hydrological events, including floods. These models can predict a variety of flooding scenarios, but since they often require heavy computational power, we cannot use them for short-time flood prediction [5]. Moreover, they have failed in forecasting floods in several cases, including the 2010 Queensland flood [6]. In recent years, machine learning (ML) algorithms have achieved unprecedented success in many tasks, ranging from object detection to speech recognition [7]. Machine learning models can extract useful patterns from data and formulate their non-linearity without requiring expert knowledge. This has encouraged many researchers to apply ML algorithms for predicting hydrological events [8].

1.2 A brief introduction to machine learning

Most relevant studies use machine learning for either predicting the flow rate of water or directly estimating the danger of floods. A machine learning model consists of the following components: (1) Train/test split (2) Pre-processing (3) Feature extraction (4) Machine Learning model (5) Evaluation [9].

- **Train/Test split**: A machine learning model is comprised of several parameters that are tuned to extract useful patterns from data for our designated task. To be able to evaluate the performance of a model, we split the dataset to train and test subsets. We use the training subset to tune the model's parameters, i.e., training the model, and evaluating its performance on the test set. It is noteworthy to

mention that the model does not encounter the test data during training.

- **Preprocessing:** Commonly, the datasets that we use for flood prediction are the outputs of sensors that are noisy. We need to remove this noise from data as it may degrade the performance of our model. In addition, we might need to modify the original data to reduce the computational cost of the algorithm, e.g., down sampling the signals. These modifications are called pre-processing and are essential for efficiently training a model.
- **Feature extraction:** Most machine learning algorithms cannot extract patterns from raw time-series data. To overcome this challenge, we need to extract useful properties, also known as features, from our raw data. These features can describe either the morphological, statistical, or any other property of a signal. Feature extraction is an important step that usually requires expert knowledge of data. It is worth mentioning that some recent ML models, such as deep neural networks, can automatically extract useful features from data and eliminate the need for this step. In most cases, after extracting features from raw data, we need to reduce their number. This step is called feature reduction and is sometimes necessary for reducing the computational cost and preventing the model from overfitting.
- **Training the model:** We feed the extracted features from the training set to our machine learning model in order to tune it. Several machine learning models can be used for flood prediction, which will be discussed in more detail in section 2.1. Each of these algorithms has its advantages and disadvantages. However, most recent studies have turned their attention to neural networks more than other ML models. A machine learning model tries to predict an output given a set of observations, i.e. training data. The output can be a discrete variable, such as whether or not the flooding occurs, or a continuous variable, such as the flow rate of water.
- **Evaluation:** Finally, to assess a proposed framework, we should evaluate its performance on the test set. Again, we emphasize that the model should not encounter the test data prior to the end of the training phase. In order to quantify the performance of a model, several metrics, such as accuracy or RMSE error, can be used depending on the continuity of the model's output.

1.3 Aim and scope of the paper

Several studies have demonstrated the ability of machine learning models in flood forecasting. The purpose of this report is to review the application of machine learning in this field. In addition, we will identify and discuss the challenges and problems that machine learning models face for flood forecasting. In the end, we offer a discussion of several state-of-the-art ML approaches that can be suitable for the task of flood prediction. The remainder of the paper is organized as follows: Section 2 briefly reviews commonly used machine learning methods in flood prediction and presents several case studies. Section 3 discusses the existing challenges in applying machine learning algorithms in flood prediction and offers new ways for using machine learning in flood forecasting. Finally, section 4 provides some concluding remarks.

2. Literature review

As earlier mentioned, several papers have successfully applied machine learning algorithms for flood forecasting. This section gives a brief review of commonly used machine learning algorithms, and then we present several case studies.

2.1 Commonly Used Machine Learning Models

In this section, we will review some of the most commonly used ML algorithms in flood prediction. It is noteworthy to mention that machine learning algorithms are not limited to the following models.

2.1.1 Support Vector Machines (SVMs)

Support Vector Machines (SVMs) are a class of machine learning models used for classification, regression, and anomaly detection. SVM is based on the theory of statistical learning and was first proposed by Cortes and Vapnik 1995. The goal of SVM is to find a vector, or equivalently a hyperplane in higher-dimensional space, that can separate data points that belong to different classes with the biggest margin. To better understand this idea, an illustration of the SVM algorithm is provided in Figure 1. Our goal is to find a hyperplane that can separate the two classes. As seen in the left figure, there is an infinite number of possible hyperplanes. The SVM algorithm tries to find the hyperplane that has the biggest margin γ from both classes (right figure).

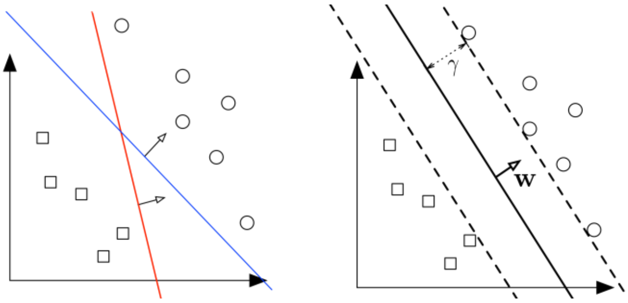


Figure 1. Left: Two possible hyperplanes that can separate data classes. Right: The supporting hyperplane that yields the maximum margin from both classes. (Figure taken from Cornell university machine learning course)

2.1.2 Decision Tree (DT) and Random Forest (RF)

A Decision Tree (DT) is a machine learning model that predicts the output using simple decision rules that learn from data. DTs can be used for predicting both continuous and discrete outputs [10]. Decision trees are simple to interpret and can be used for fast prediction. On the other hand, they are vulnerable to overfitting and cannot generalize well. To address this issue, researchers have proposed using several decision trees simultaneously for predicting the output. The resulting algorithm is referred to as random forest (RF) [11]. In RF, each individual decision tree outputs a prediction for the target variable. By assembling these predictions, we can obtain a more robust model with less overfitting.

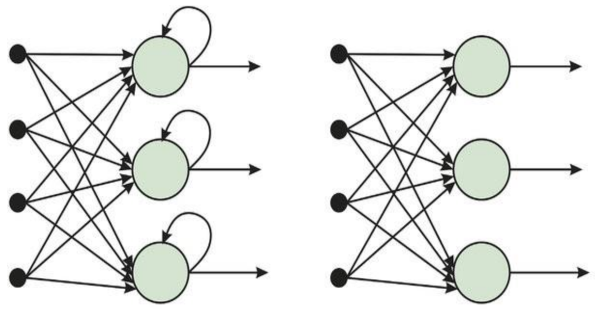
2.1.3 Feed-Forward Neural Networks (FFNN)

Deep learning is a subfield of machine learning that has gained popularity in recent years due to its capability to learn complex and nonlinear patterns from raw data [9]. The idea behind deep learning models is to mimic the function of the human nervous system [12]. Because of this, deep learning models are also sometimes referred to as neural networks. Feed-Forward Neural Networks, also known as Multilayer Perceptrons (MLPs), are quintessential deep learning algorithms. An FFNN consists of an input layer, one or more hidden layer(s), and one output layer. Each layer is comprised of several neurons. A neuron is a linear function, followed by a nonlinearity. During the training phase, the parameters of neurons, i.e. their weights, are tuned using a procedure called back-propagation, combined with optimization algorithms such as stochastic gradient descent (SGD) [13]. In

feedforward NNs, the information flows from the input to the output, i.e. there is no feedback between the hidden layers.

2.1.4 Recurrent Neural Networks (RNNs)

Despite their unique properties, feedforward neural networks fail to capture the dependencies in sequential data. Recurrent Neural Networks (RNNs) can solve this problem by introducing connections between the hidden units. The idea behind these feedback connections is that they add a memory to the system that stores valuable information from the earlier training steps [14]. Figure 2 depicts the architecture of an RNN against a feedforward neural network. The main difference between these two architectures is the recurrent connections in the hidden layers of the RNN. In RNN, the output of a hidden layer is a function of its input, as well as its state in the previous time step. This helps the network to remember the important information from the past.



(a)Recurrent Neural Network (b) Feed-Forward Neural Network

Figure 2. (a) Architecture of a simple RNN (b) Architecture of a FeedForward Neural Network Figure taken from [15].

Parameters of RNNs are tuned by a procedure called backpropagation through time (BPTT). RNNs are suitable for sequential data, such as time-series signals and natural language, but they face problems such as gradient vanishing. Some modifications of RNN, such as Long Short Term Memory (LSTM) and Gated ReLU Unit (GRU), are proposed to overcome this problem [16].

2.2 Examples of machine learning application in flood forecasting

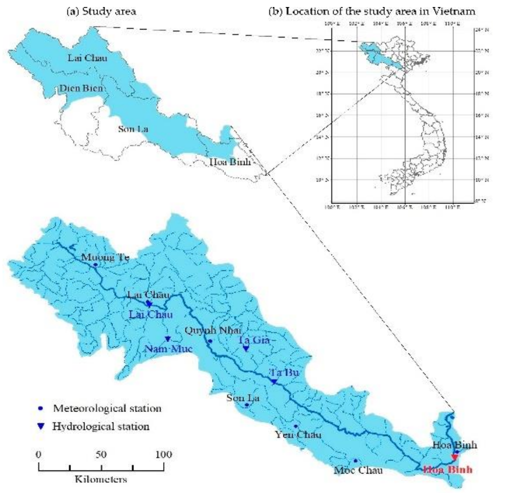
In this section, we briefly go over three recent papers that have utilized ML models for flood forecasting.

2.2.1 Case Study 1: *Flood Prediction in the River Nile, Sudan Using Neural Network*

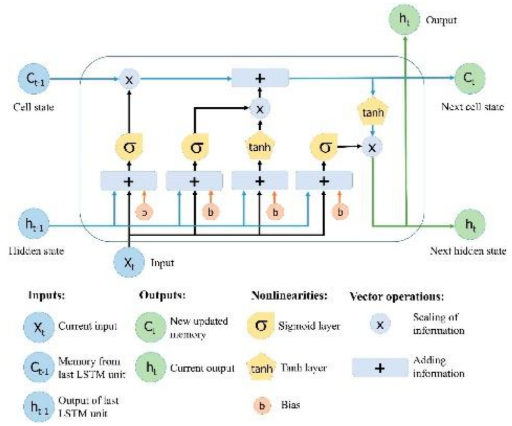
This study [17] proposed a neural network for predicting the flow rate of water at Dongola station in the Nile River, Sudan. The model's input is upstream flow data, and it tries to predict the downstream flow rate with one day lead. The proposed model achieved an accuracy of around 97% in detecting flood hazards. The significance of this study was that the author could achieve this performance by only using one variable, i.e., upstream flow data. Mathematical models that simulate the hydrodynamic process of the water's flow can also achieve good accuracy, but they often require large datasets. This shows an advantage of NNs in flood forecasting over conventional models.

2.2.2 Case Study 2: Flood Forecasting Using LSTM in Da River Basin, Vietnam

In this study [18], the authors applied an LSTM model for forecasting floods on the Da river, which is one of the largest river basins in Vietnam. Their model tries to forecast the river's flow using rainfall and flowrate data acquired at hydrological stations of the study area.



(a) Map of The Study Location



(b) Overview of The LSTM Model

Figure 3. Location of study and LSTM architecture of the paper Le et al. [18]

Da river originates from Yunnan province in China and flows through high mountains in Vietnam. It has an abundant flow which makes it a suitable source for hydroelectric power. Due to the existence of dams upstream of the Da river, forecasting its downstream flow is a challenging task. The paper attempted to predict the downstream flow with one, two, and three days of lead time using the stations in the Hoa Binh area. The location of the study area and stations are illustrated in Figure 3 (a). An LSTM is a special kind of RNN that can learn long-term dependencies of the data. The structure of an LSTM model is shown in Figure 3 (b). The network consists of memory blocks called cells. These cells are organized in a way that the network can remember long-term dependencies in the data. To evaluate the performance of their method, they used the NSE and RMSE criteria and have shown that their model can accurately predict the flow rate with one, two, and three days lead. As expected, the accuracy of their prediction with one day lead is better than other scenarios. The results of their study are summarized in Table 1.

Table 1. Results of the LSTM model for flow rate prediction

Predict for	RMSE (m3/s)	NSE (%)	Predicted Peak	Observed Peak	**Relative Error** (%)
One day	152.4	99.1	9340	10000	6.6
Two days	360.7	94.9	8477	10000	15.2
Three days	571.4	87.2	7181	10000	28.2

It is worthy of mentioning that the LSTM model does not use any expert knowledge but instead learns the physical principles such as principles of mass conservation or momentum conservation during training from the input data. This study showed that LSTMs could be reliable tools for predicting the flow rate and flood risk with fair accuracy.

2.2.3 Case Study 3: Urban Flood Prediction Using NNs with Data Augmentation

This study [19] uses a deep neural network to estimate the total accumulative overflow. Since the amount of available heavy rainfall and actual flood is limited, they used data augmentation techniques to train the model with more data. Their study area was the Samseong-dong district in Seoul, Korea. Due to its low land and insufficient drainage system, this area experienced an extreme urban flood in 2010 and 2011. Since it is a metropolitan area, modelling the flood flow requires complex calculations. Statistical properties of each rainfall event were used as the input. They compared the prediction of their model with the results of computer simulation. They confirmed that the performance of their model is improved when they use data augmentations. Lack of historical data is an important challenge in most flood prediction problems. On the other hand, the performance of machine learning models heavily depends on the amount of available training data. The significance of this study is that the authors proposed using data augmentation and confirmed that it could improve the model's performance.

3. Discussion

Machine learning models have shown promising performance in forecasting floods. Existing algorithms can predict the danger of floods with fair accuracy and lead time. However, there are still several challenges that ML models face that should be taken into account. In addition, proposed methods for flood forecasting are mainly based on a time-series prediction algorithm. Time-series prediction is an active area of research in machine learning, and every year, new methods are introduced that can outperform the preceding algorithms [20]. These state-of-the-art algorithms can also improve the performance of flood prediction models. In the following sections, we first critically evaluate the usage of ML models in flood prediction and the challenges that this research area faces. Then, we propose several new avenues for improving the existing models.

3.1 Existing Challenges

3.1.1 Data Availability

As earlier mentioned, the success of an ML model depends on the quality and amount of available data. Depending on the economic development of a country, different datasets are available for training a model. With more available data, the performance of an ML model improves. We should note that a study that reports an accurate performance on a given dataset may fail in other cases where we do not have access to adequate data. With that being said, the amount of required data for an ML model is still less in comparison to physically-based models. Another important consideration for using ML models is their vulnerability to noise. In practical scenarios, our data is often noisy, and this can significantly degrade the performance of most machine learning models [21]. In these cases, careful pre-processing is required to ensure that the input of the ML model is clean. Finally, an important issue of machine learning models is a phenomenon called the curse of dimensionality [22]. This problem occurs when the number of samples is limited compared to the number of features. Since most ML models in the literature use a limited number of input features for flood prediction, this problem is not reported in this field. However, if a model is designed to work with more features, it should also include a large number of samples to avoid this problem.

3.1.2 Lack of Benchmark for Comparison

As earlier mentioned, past studies report the performance of their model on different datasets. Therefore, it is hard to judge the accuracy of different algorithms for flood prediction. To solve this problem, all proposed algorithms should be compared together on one unified dataset. For example, In computer vision, there are several benchmark datasets, such as CIFAR-10 and CIFAR-100, which include different images from various objects, ranging from animals to vehicles [7]. When a team of researchers proposes a new object detection algorithm, they commonly report the performance of their model on these datasets. This way, we can compare these algorithms with their older counterparts to see how much improvement they can make. Another important shortcoming of previous studies is that they do not compare their proposed model with other state-of-the-art algorithms. In most standard machine learning papers, the performance of the proposed approach is compared to several other competing models, also called baselines. However, in most flood prediction papers, the proposed ML approach is either not compared to other ML models or is just compared to simple statistical or physically-based models. Due to the lack of a benchmark dataset, an efficient model which is trained on a complex dataset can achieve a worse performance in comparison to a simple model that is trained on easily predictable data; however, if there exists a standard benchmark dataset with a variety of samples from different conditions, we can better evaluate the performance of an algorithm as well as the scenarios that it can work well or fail.

3.1.3 How much can we trust machine learning models?

An important question to ask is how much we can trust machine learning algorithms in predicting flood danger. To answer this question, we should take another look at the process of training an ML model. A machine learning model learns to extract patterns from the input data. Therefore, its performance depends on the quality of input data that we provide it with. If the data is noisy or has no patterns, we cannot expect the ML model to yield acceptable performance. Another important consideration is that a model that is trained on a specific dataset might not yield the same performance on other datasets. This is sensible since each geographical area has its own properties, and we cannot expect a model to forecast all the scenarios that it has not encountered. Overall, we should have a realistic expectation from machine learning models. They are as good as the data we provide them with, and we cannot expect them to be helpful in situations where our data has no specific pattern or is too noisy. With that being said, ML models are still powerful tools for designing flood alarm systems. They can rapidly

extract meaningful information from large datasets, a task which is almost impossible for human experts to handle manually. But still, we should keep in mind that they have limitations and cannot always yield the same performance.

3.2 Future of Machine Learning in Flood Prediction

This section proposes some novel ways for applying machine learning models in flood prediction that can be further explored in future studies.

3.2.1 Incorporating Expert Knowledge

Many previous studies have mentioned that an advantage of machine learning models over other flood prediction methods is that they do not require any expert knowledge and can export important information solely from the input data. However, an immense amount of research has been carried out by hydrologists to understand different aspects of the flood. This information can enhance the performance of an ML model and help us to design physically feasible models. To this date, most studies have focused on designing purely data-driven models. However, incorporating existing knowledge in these models can be a new research track that can lead to more efficient models.

3.2.2 Hybrid Models

In section 2.1, several machine learning models have been reviewed. Each of these models had its advantages and disadvantages. A recent trend in machine learning research is developing hybrid models that use two or more models for carrying out a specific task. The output of these models is then combined to produce the final model's output. The hybrid approaches can be used to combine the advantages of two or more models and design a more accurate model. So far, several studies have used hybrid models for flood forecasting, but this research is still not fully explored and has a lot of potential for performing more experiments [23].

3.2.3 Using State-of-the-art Time Series Prediction Models

As earlier mentioned, most ML-based flood forecasting models are inspired by recent developments in time-series prediction. Therefore, it would be worthy of looking at some of the current models that are considered to be state-of-the-art in this field. Transformers are one of these models that have recently gained huge success in different tasks [24]. Like RNNs, they are suitable for handling sequential data. They have also been successfully used for forecasting weather-related data [25].

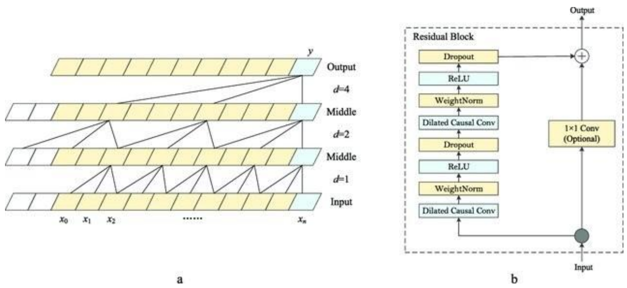


Figure 4. (a) Overview of a 1D Convolutional Network with Dilated Convolution (b) Architecture of a TCN block (Figure taken from [26])

Another state-of-the-art algorithm is called Temporal Convolutional Network (TCN). Convolutional Neural Networks (CNNs) have achieved massive success in computer vision, and TCN models are very similar to them but are modified for 1D outputs, such as time-series. TCNs use the idea of dilated convolution, which helps capture information from different time scales of a signal. They have shown to perform well on weather forecasting problems and are considered suitable candidates for flood prediction (Hewage et al. 2020). Figure 4 depicts the architecture of a TCN model.

4. Conclusions

In this paper, we have reviewed the application of machine learning models in flood prediction. We first gave an overview of the problem, as well as some commonly used machine learning algorithms. Then, we went over some case studies that showcase the ability of ML-based models in flood prediction. Finally, we discussed the existing challenges in applying ML models and proposed some possible directions for the future of ML in flood prediction research. Overall, machine learning has become a popular field in recent years and has replaced older models in many fields. It is evident that ML is gaining a lot of attention from hydrologists for flood prediction. If they are used in an appropriate fashion, they can be a powerful tool for generating flood alarms and for alleviating the damages.

Ethical issue

The authors are aware of and comply with best practices in publication ethics, specifically with regard to authorship (avoidance of guest authorship), dual submission, manipulation of figures, competing interests, and compliance with policies on research ethics. The authors adhere to publication requirements that submitted work is original and has not been published elsewhere in any language.

Data availability statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflict of interest

The authors declare no potential conflict of interest.

References

[1] Floods n.d. https://www.who.int/health-topics/floods#tab=tab_1 (accessed March 22, 2022).

[2] Floods n.d. https://www.publicsafety.gc.ca/cnt/mrgnc-mngmnt/ntrl-hzrds/fld-en.aspx (accessed March 22, 2022).

[3] Ohl CA, Tapsell S. Flooding and human health. BMJ 2000;321:1167–8. https://doi.org/10.1136/BMJ.321.7270.1167.

[4] Basha EA, Ravela S, Rus D. Model-based monitoring for early warning flood detection. SenSys'08 - Proc 6th ACM Conf Embed Networked Sens Syst 2008:295–308. https://doi.org/10.1145/1460412.1460442.

[5] Ji J, Choi C, Yu M, Yi J. Comparison Of A Data-driven Model And APhysical Model For Flood Forecasting. WIT Trans Ecol Environ 2012;159:133–42. https://doi.org/10.2495/FRIAR120111.

[6] Van Den Honert RC, Mcaneney J. The 2011 Brisbane Floods: Causes, Impacts and Implications. Water 2011, Vol 3, Pages 1149-1173 2011;3:1149–73. https://doi.org/10.3390/W3041149.

[7] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks n.d.

[8] Mosavi A, Ozturk P, Chau KW. Flood Prediction Using Machine Learning Models: Literature Review. Water 2018, Vol 10, Page 1536 2018;10:1536. https://doi.org/10.3390/W10111536.

[9] Deep Learning | The MIT Press n.d. https://mitpress.mit.edu/books/deep-learning (accessed March 22, 2022).

[10] Tehrany MS, Pradhan B, Jebur MN. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. J Hydrol 2013;504:69–79. https://doi.org/10.1016/J.JHYDROL.2013.09.034.
[11] Breiman L. Random Forests. Mach Learn 2001 451 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
[12] Mcculloch WS, Pitts W. A logical calculus of ideas immanent in nervous activity. Bull Mothemnticnl Biol 1990;52:99–115.
[13] Ruder S. An overview of gradient descent optimization algorithms 2016. https://doi.org/10.48550/arxiv.1609.04747.
[14] Sherstinsky A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Phys D Nonlinear Phenom 2020;404:132306. https://doi.org/10.1016/J.PHYSD.2019.132306.
[15] Eliasy A, Przychodzen J. The role of AI in capital structure to enhance corporate funding strategies. Array 2020;6:100017. https://doi.org/10.1016/J.ARRAY.2020.100017.
[16] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput 1997;9:1735–80. https://doi.org/10.1162/NECO.1997.9.8.1735.
[17] Elsafi SH. Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. Alexandria Eng J 2014;53:655–62. https://doi.org/10.1016/J.AEJ.2014.06.010.
[18] Le XH, Ho HV, Lee G, Jung S. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. Water 2019, Vol 11, Page 1387 2019;11:1387. https://doi.org/10.3390/W11071387.
[19] Kim H Il, Han KY. Urban Flood Prediction Using Deep Neural Network with Data Augmentation. Water 2020, Vol 12, Page 899 2020;12:899. https://doi.org/10.3390/W12030899.
[20] Fu TC. A review on time series data mining. Eng Appl Artif Intell 2011;24:164–81. https://doi.org/10.1016/J.ENGAPPAI.2010.09.007.
[21] Liu M, Liu S, Su H, Cao K, Zhu J. Analyzing the Noise Robustness of Deep Neural Networks. 2018 IEEE Conf Vis Anal Sci Technol VAST 2018 - Proc 2018:60–71. https://doi.org/10.1109/VAST.2018.8802509.
[22] Keogh E, Mueen A. Curse of Dimensionality. Encycl Mach Learn Data Min 2017:314–5. https://doi.org/10.1007/978-1-4899-7687-1_192.
[23] Li X, Willems P. A Hybrid Model for Fast and Probabilistic Urban Pluvial Flood Prediction. Water Resour Res 2020;56:e2019WR025128. https://doi.org/10.1029/2019WR025128.
[24] Vaswani A, Brain G, Shazeer N, Parmar N, Uszkoreit J, Jones L, et al. Attention Is All You Need n.d. https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (accessed March 22, 2022).
[25] Wu N, Green B, Ben X, O'Banion S. Deep Transformer Models for Time Series Forecasting: The Influenza Prevalence Case 2020. https://doi.org/10.48550/arxiv.2001.08317.
[26] Bai S, Kolter JZ, Koltun V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling 2018. https://doi.org/10.48550/arxiv.1803.01271.